High-Resolution Genetic Mapping of Complex Traits from a Combined Analysis of F2 and Advanced Intercross Mice

Clarissa C. Parker,*1 Peter Carbonetto,† Greta Sokoloff,† Yeonhee J. Park,† Mark Abney,† and Abraham A. Palmer†,‡

*Department of Psychology and Program in Neuroscience, Middlebury College, Middlebury, Vermont 05753, and †Department of Human Genetics and ‡Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, Chicago, Illinois 60637

ABSTRACT Genetic influences on anxiety disorders are well documented; however, the specific genes underlying these disorders remain largely unknown. To identify quantitative trait loci (QTL) for conditioned fear and open field behavior, we used an F2 intercross (n = 490) and a 34th-generation advanced intercross line (AIL) (n = 687) from the LG/J and SM/J inbred mouse strains. The F2 provided strong support for several QTL, but within wide chromosomal regions. The AIL yielded much narrower QTL, but the results were less statistically significant, despite the larger number of mice. Simultaneous analysis of the F2 and AIL provided strong support for QTL and within much narrower regions. We used a linear mixed-model approach, implemented in the program QTLRel, to correct for possible confounding due to familial relatedness. Because we recorded the full pedigree, we were able to empirically compare two ways of accounting for relatedness: using the pedigree to estimate kinship coefficients and using genetic marker estimates of “realized relatedness.” QTL mapping using the marker-based estimates yielded more support for QTL, but only when we excluded the chromosome being scanned from the marker-based relatedness estimates. We used a forward model selection procedure to assess evidence for multiple QTL on the same chromosome. Overall, we identified 12 significant loci for behaviors in the open field and 12 significant loci for conditioned fear behaviors. Our approach implements multiple advances to integrated analysis of F2 and AILs that provide both power and precision, while maintaining the advantages of using only two inbred strains to map QTL.

ANXIETY disorders are among the most prevalent psychiatric disorders in the world; in the United States, they affect the lives of ~18% of the adult population (Demyttenaere et al. 2004; Kessler et al. 2005a,b). Many of these debilitating illnesses can be characterized by exaggerations of normal and adaptive emotional response to fearful or stressful events (Mahan and Ressler 2012). Twin and family studies support a genetic basis for anxiety disorders, but attempts to identify the underlying genetic substrates have been disappointing—to date, genome-wide association studies (GWAS) have not reliably replicated candidate genes associated with anxiety disorders (Hettema et al. 2011). As a result, we have little knowledge of the specific genes that underlie these disorders. Genetic loci relevant to these disorders may be difficult to map via GWAS because anxiety disorders are only modestly heritable and, like many psychiatric conditions, are expected to be highly polygenic—that is, modulated by a large number of genetic factors with individually small effects (Sullivan et al. 2012). Therefore, genetic contributions to anxiety are likely to be difficult to distinguish from correlations that occur by chance alone (Flint 2011; Parker and Palmer 2011; Flint and Eskin 2012).

While the full spectrum of any human psychiatric disorder can never be fully recapitulated in a single mouse model, there is substantial behavioral, genetic, and neuroanatomical conservation between humans and mice. When broken down into individual components, many of the symptoms of anxiety disorders can be modeled in mice. Thus, translational mouse models can provide a powerful strategy for understanding the genetic and biological underpinnings of
the acquisition of fear, as well as the etiologic processes related to anxiety (Kaloufi et al. 2007; Flint and Shifman 2008; Hovatta and Barlow 2008). Animal models provide similarly strong models for comorbidity of traits, like evidence for a shared genetic substrate in anxiety and fear. For example, selective breeding paradigms in mice and rats have shown that selection of anxiety-like behavior also selects for differences in fear and vice versa (Ponder et al. 2007a; López-Aumatell et al. 2009).

Reverse genetic approaches using genetically modified mice have been important for testing hypotheses about specific genes relevant to fear and anxiety, but they have tended to focus on the “usual suspects” underlying anxiety disorders. Alternatively, forward genetic approaches in mice have been developed to measure phenotypes and identify the underlying sources of standing genetic variability in these phenotypes without prior hypotheses. However, forward genetic approaches have been less successful at identifying relevant genes. This may be because forward genetic studies in mice have traditionally used recombinant inbred (RI) lines, backcrosses (BC), or F2 intercrosses to identify quantitative trait loci (QTL)—in these experimental crosses, we can have high statistical power to detect genetic loci, but poor mapping resolution due to limited recombination (Cheng et al. 2010; Flint 2011; Parker and Palmer 2011).

To better pinpoint candidate genes and genetic loci that might influence anxiety, we mapped QTL in a combined F2 intercross and an F34 advanced intercross line (AIL). An AIL is created by successive generations of pseudorandom mating after the F2 generation. Each additional generation leads to the accumulation of new recombinations, which allows for more precise mapping due to a breakdown in linkage disequilibrium. We show that our analysis not only yields strong support for several QTL in anxiety-related phenotypes, but also highlights a narrower set of candidate genes than previous studies of these phenotypes.

AILs have been employed in several previous studies to successfully map QTL for complex traits in mice, including locomotor activity (Cheng et al. 2010), muscle weight (Lionikas et al. 2010), red blood cell characteristics (Bartnikas et al. 2012), body weight (Parker et al. 2011), methamphetamine sensitivity (Parker et al. 2012a), prepulse inhibition (Samocha et al. 2010), and the conditioned fear phenotypes studied in this article (Parker et al. 2012b). This study makes several key contributions over previous work in this area. First, we demonstrate that the combination of an F2 intercross and a 34th-generation AIL allows us to map more QTL for conditioned fear and anxiety-related phenotypes and at a greater precision than our earlier study of an 8th-generation AIL (Parker et al. 2012b). Second, unlike previous studies, we use SNP data to account for the varying levels of genetic sharing that may confound detection of QTL. This is particularly relevant because assessing evidence for QTL in an AIL population is more challenging than in traditional designs such as F2 intercrosses because varying levels of genetic sharing among individuals in the AIL can confound tests for association (Cheng et al. 2010). Third, we demonstrate the benefits of using marker data in place of a pedigree to infer familial relationships in an AIL. We find that the linear mixed model with marker-based relatedness estimates yields greater support for QTL compared to using the pedigree to estimate genetic sharing, a finding that is consistent with a recent comparison of these approaches in simulated populations (Cheng and Palmer 2013). Finally, we show that a simple permutation test that assumes independence of the samples—that is, the permutation test ignores the varying levels of genetic sharing—provides an adequate way to assess significance of QTL in our AIL population.

Materials and Methods

Animals and housing

Subjects consisted of 487 F2 mice (249 males and 241 females) derived from a cross between LG/J and SM/J inbred strains obtained from The Jackson Laboratories (Bar Harbor, ME) and 687 F34 LG/J × SM/J AIL mice (353 males and 334 females) derived from F33 breeders obtained from the laboratory of James Cheverud (Washington University, St. Louis). Three mice were later removed from the F2 cohort because a high proportion of their markers were not reliably genotyped. The colony was maintained on a 12:12-hr light:dark cycle with lights on at 0630 hr. Mice were housed in clear acrylic cages with corn-cob bedding in same-sex groups of 2–5 mice with food and water available ad libitum. The full pedigree of all AILs was recorded so that the ancestry of each mouse could be traced back to the inbred founders.

To produce F34 mice from the F33 breeders, we mated the F33 mice such that no mating pairs shared a common grandparent. (In subsequent generations, which are not described in this article, we selected pairs of mice for breeding that produced the smallest possible inbreeding coefficients in the offspring. This was achieved using R code written by Andrew Skol, which we have made available at http://github.com/pcaibo/breedail.)

Behavioral testing

Behavioral testing in all mice was always conducted during the light phase, between 0800 and 1700 hr. We waited for the mice to acclimatize to the testing room for at least 30 min before beginning the tests. Mice were ~2–3 months of age on the first day of behavioral testing (F2 range was 53–71 days; AIL range was 50–76 days). All mice went through an identical testing sequence: first, we measured open field (OF) behavior as part of a locomotor testing paradigm (Bryant et al. 2012); 1 week later, we began the conditioned fear (CF) paradigm. All experiments were performed in accordance with the National Institutes of Health guidelines for care and use of laboratory animals. Experimental procedures were approved by the University of Chicago’s Institutional Animal Care and Use Committee.

OF

Our procedures for OF testing have been explained in detail in previous publications (Palmer et al. 2005; Sokoloff et al. 2012).
Briefly, after 30 min of acclimation to the testing room, mice were removed from the home cage and placed into individual holding cages for 5 min, and then they were weighed individually, injected i.p. with physiological saline (0.01 ml/g body weight), and immediately placed in the center of the OF (AccuScan Instruments, Columbus, OH). Each OF was housed inside a sound-attenuating environmental chamber (AccuScan Instruments) with overhead lighting providing illumination (~80 lux) and a fan in the rear wall providing ventilation. Mice were tested in the OF arena (40 × 40 × 30 cm) for 30 min during the locomotor test. The first 10 min were used to measure OF behavior. We measured three phenotypes (Versamax, AccuScan Instruments): (1) distance traveled (centimeters) in the periphery (width: 10 cm), (2) distance traveled (centimeters) in the center (20 × 20 cm), and (3) proportion of time spent in the center of the arena. We mapped QTL for these three phenotypes. After testing, mice were returned to their home cage, and the field was cleaned with 10% isopropanol before the next mouse was tested. At the end of testing, mice were returned to the vivarium. Distance traveled from this same session was also analyzed in a previous article (Cheng et al. 2010).

**CF**

CF procedures were identical to those described previously (Ponder et al. 2007a). Briefly, mice were tested in standard CF chambers (29 × 19 × 25 cm) housed within sound-attenuating chambers (Med Associates, St. Albans, VT). Lights in each chamber provided dim illumination (~3 lux), and fans provided a low level of masking background noise. Chambers were cleaned with 10% isopropanol between animals. Behavior was digitally recorded by a computer and subsequently analyzed with FreezeFrame software (Actimetrics, Evanston, IL).

Testing for CF consisted of a 5-min test that occurred three times over three consecutive days. After habituating to the procedure room for 30 min in their home cages, mice were transferred to the CF chambers in individual holding cages. On day 1, baseline freezing (“pretraining freezing”) was measured beginning 30 sec after mice were placed in the test chambers and ending 150 sec later. After the pretraining period, mice were exposed twice to the conditioned stimulus (CS), a 30-sec tone (85 dB, 3 kHz) that coterminated with the unconditioned stimulus (US), a 2-sec, 0.5-mA foot shock delivered through the stainless steel floor grid. After each CS–US pairing, there was a 30-sec period in which no stimuli were delivered to the subject.

Test day 2 began exactly 24 hr after the start of test day 1. The testing environment was identical to that of day 1, except that neither tones nor shocks were presented. Proportion of time freezing in response to the test chamber (“% freezing to context”) was measured over the same time period as pretraining freezing (30–180 sec). We chose this time period to permit immediate comparison to the pretraining freezing scores on day 1 and to avoid measuring freezing behavior during the latter part of the trial in which the mice might have anticipated shocks based on tests from previous days.

DNA from the F2 mice was extracted by LGC Genomics (Hoddesdon, Hertfordshire, UK; formerly KBiosciences). F2 genotypes were called using KASPar, a fluorescence-based PCR assay (LGC Genomics), at 162 evenly spaced markers on autosomal and X chromosomes. These markers are a subset of the 1638 SNPs suggested by Petkov et al. (2004) for QTL mapping in mouse strains. AIL genotypes at 4601 SNPs on autosomal and X chromosomes were ascertained using an Illumina Infinium Platform (iSelect) custom genotyping array (http://www.illumina.com), as described previously (Cheng et al. 2010). To avoid having to impute unascertained genotypes in the AIL mice, we discarded 7 of the 162 SNPs genotyped in the F2 cohort that were not genotyped in the F34 samples. Of the final set of 4601 candidate SNPs, 4535 (98.6%) correspond to SNPs in the dbSNP reference database (Sherry et al. 2001). A list of SNPs used in the present study is available at the Mouse Phenome Database (http://phenome.jax.org) under project name “Chicago1” (http://phenome.jax.org/db/q?rtn=projects/projdet&reqprojid=316).

The Chicago1 data set includes ~8200 SNPs; of these, 4601 are polymorphic between LG/J and SM/J mice. All SNP identifiers and locations of the SNPs reported in this article are based on release 37 (July 2007) of the NCBI Mouse Genome Assembly. Average SNP allele frequencies in the F2 and F34 cohorts were 0.50 and 0.51, respectively, indicating little unintended selection or genetic drift. As expected, we observed greater variation in allele frequencies in the F34 cross; SNP allele frequencies in F2 mice ranged from 0.45 to 0.54, whereas allele frequencies in F34 mice ranged from 0.12 to 0.90. (See Cheng et al. 2010 for further discussion of genetic drift in the F34 AIL.)

Genotypes called with lower confidence were removed and treated as missing. In the F2 mice, <1% of the genotypes were treated as missing. At most 9% of genotypes were missing for any single SNP and at most 26% in any F2 mouse (after removing the three mice with a large fraction of poor-quality genotype calls). In the F34 mice, 1% of the genotypes were treated as missing overall; at most 1.2% of genotypes were missing for any single SNP and at most 5% in any F34 mouse.
**Genotype imputation**

To analyze the combined data, we estimated the probabilities of missing genotypes and the genotypes of SNPs that were called in the F34 mice and not called in the F2 mice. Probabilities of missing or unavailable genotypes in the F2 and F34 crosses were estimated based on correlation patterns with available genotypes, using recombination models for autosomal chromosomes in advanced intercrosses (Darvasi and Soller 1995). Genotype probabilities were calculated separately in F2 and F34 crosses as they exhibited different patterns of recombination. Even though only a small fraction of the full set of SNPs were genotyped in the F2 mice (155 of 4601), this small panel of SNPs was mostly sufficient to accurately estimate genotypes at other SNPs, since F2 chromosomes experience little recombination, and most alleles occur at high frequencies. In the F2 mice, 85% of genotypes were imputed with high confidence (maximum genotype probability >0.9). Only a small fraction of SNPs were imputed less accurately, presumably because they were not in sufficient linkage disequilibrium with a SNP genotyped in the F2 mice; for 109 SNPs (2.4%), <20% of the genotypes in the F2 mice were estimated with high confidence (maximum genotype probability >0.9). Since we accounted for genotype uncertainty in the QTL mapping, the contribution of each sample to the association signal was appropriately adjusted to reflect the confidence of the genotype estimates. We used the Haley–Knott approximation (Broman and Sen 2009) to account for uncertainty in missing genotypes. Note that we did not perform interval mapping to assess support for QTL between markers.

Estimation of missing genotypes requires genetic distance estimates at all markers. We used the genetic distance estimates from the reference panel described in Cox et al. (2009). These estimates were retrieved from the Mouse Map Converter hosted at Jackson Laboratories (http://cgdjax.org/mousemapconverter). We checked these genetic distances against intermarker distances estimated from our F2 genotypes. Estimates were obtained using the Lander–Green method (Lander and Green 1987), as implemented in R/qtl (Broman and Sen 2009). Intermarker distances on the X chromosome were estimated using females – where a pair of randomly chosen alleles from individuals i and j at the same autosomal locus are identical by descent (IBD); and variables to be fitted to the data.

QTLRel uses a two-step procedure to fit the mixed model to the data: first, the model parameters are estimated assuming that no markers have an effect on the phenotype (null hypothesis); second, for a given marker, the additive and dominance effects are estimated while the other model components are fixed to their values obtained from the first step, up to a scaling factor that is also estimated in this step. This is similar to the strategy used in EMMAX (Kang et al. 2010). [More recently, efficient methods have been developed to integrate model fitting into a single step, which can sometimes lead to improved power to detect QTL (see Lippert et al. 2011; Zhou and Stephens 2012)]). Once the model-fitting steps are completed, QTLRel uses the parameter estimates to compute the log-likelihood-ratio test statistic. For each SNP, we reported support for a genotype–phenotype association using the (base 10) logarithm of the likelihood ratio, commonly called the LOD score.

Since F2 crosses are approximately equally related to each other, QTL mapping that does not account for familial relationships should yield similar results to those of QTLRel. To verify this, we compared our results in the F2 sample using QTLRel against results from a linear regression model that did not include the polygenic component, implemented in R/qtl (Broman and Sen 2009). Conversely, ignoring relatedness in the F34 and combined cohorts is expected to yield very different QTL mapping results (Cheng et al. 2010).

**QTL mapping**

**Linear mixed model for QTL mapping:** Individuals in an F2 cross are full sibs that share roughly the same amount of the genome, so it is common to map QTL using a simple linear regression approach that ignores familial relatedness. By contrast, genetic sharing can vary considerably in an AIL, so it is important to correct for possible confounding due to varied levels of relatedness (Abney et al. 2000; Cheng et al. 2010). We used the QTLRel mixed-model framework, which was developed for QTL mapping in AILs (Cheng et al. 2010, 2011). What makes QTLRel different from a standard linear regression is the inclusion of a “polygenic effect” that captures how correlations in the phenotype are explained by genome-wide genetic sharing. QTLRel models the phenotype as a linear combination of these variables: additive genotype (allele count), dominance genotype (0 = heterozygous, 1 = homozygous), additional covariates such as sex and coat color, the polygenic effect, and the residual. (Note that the X chromosome requires a separate treatment, as we explain below.) Fitting this model to the data involves estimating the $n \times n$ covariance matrix of the polygenic effect, where $n$ is the number of samples. While the general expression for this covariance matrix is derived in Abney et al. (2000) for analyzing quantitative traits in an inbred population, that article suggests using a simpler expression in which only the additive and dominance terms are retained. Following this suggestion, the covariance matrix entry corresponding to pairs of individuals $(i,j)$ is $\sigma_{ij} = 2\Phi_{ij}\sigma_A^2 + \Delta_{ik}\sigma_D^2$, where $\Delta_{ik}$ is the conditional probability of identity state $k$ for pair $(i,j)$, otherwise known as the “condensed coefficient of identity”; $\Phi_{ij} = \Delta_{ij,1} + (1/2)(\Delta_{ij,3} + \Delta_{ij,5} + \Delta_{ij,7}) + (1/4)\Delta_{ij,8}$ is the kinship coefficient for pair $(i,j)$—i.e., the probability that a pair of randomly chosen alleles from individuals $i$ and $j$ at the same autosomal locus are identical by descent (IBD); and $\sigma_A^2$ and $\sigma_D^2$ are parameters to be fitted to the data.

QTLRel uses a two-step procedure to fit the mixed model to the data: first, the model parameters are estimated assuming that no markers have an effect on the phenotype (null hypothesis); second, for a given marker, the additive and dominance effects are estimated while the other model components are fixed to their values obtained from the first step, up to a scaling factor that is also estimated in this step. This is similar to the strategy used in EMMAX (Kang et al. 2010). [More recently, efficient methods have been developed to integrate model fitting into a single step, which can sometimes lead to improved power to detect QTL (see Lippert et al. 2011; Zhou and Stephens 2012)]. Once the model-fitting steps are completed, QTLRel uses the parameter estimates to compute the log-likelihood-ratio test statistic. For each SNP, we reported support for a genotype–phenotype association using the (base 10) logarithm of the likelihood ratio, commonly called the LOD score.

Since F2 crosses are approximately equally related to each other, QTL mapping that does not account for familial relationships should yield similar results to those of QTLRel. To verify this, we compared our results in the F2 sample using QTLRel against results from a linear regression model that did not include the polygenic component, implemented in R/qtl (Broman and Sen 2009). Conversely, ignoring relatedness in the F34 and combined cohorts is expected to yield very different QTL mapping results (Cheng et al. 2010).
(Abney et al. 2000; Cheng et al. 2010). Specifically, the additive and dominance variance components of this covariance matrix require estimates of the kinship coefficient $\Phi_{ij}$ and identity coefficient $\Delta_{ij}$ for each pair of individuals $(i,j)$. We derived estimates of identity coefficients in two different ways, using the pedigree, and using the marker data. Algorithms to compute identity coefficients from the pedigree are far too complex for large pedigrees such as the one used here, so we applied the scalable approximation developed by Cheng et al. (2010), building on Karigl (1981). Even this approximate algorithm was not sufficiently scalable for computation of coefficients for $F_2-F_{34}$ pairs, so for each of these pairs we assigned the identity coefficients to their expected values in an inbred cross, $\Phi_{ij} = 1/2$ and $\Delta_{ij} = 1/4$.

One advantage of using the markers to estimate relatedness was that it was much easier to compute identity coefficient estimates for all pairs of individuals. In keeping with other methods that use marker-based estimates of relatedness to correct for confounding due to population stratification or familial relationships (Yu et al. 2006; Kang et al. 2008; Lippert et al. 2011; Listgarten et al. 2012; Zhou et al. 2013), we included only the additive portion of the polygenic effect. The rationale is that dominance effects typically make a much smaller contribution to the variance of a complex trait. Therefore, we needed to calculate only kinship coefficients. At a single locus, the estimate of $2\Phi_{ij}$ is simply equal to the number of alleles that share the same state between individuals (2 if both genotypes are homozygous and the same, 0 if both genotypes are homozygous and different, and 1 in all other cases). This is equivalent to the number of alleles IBD, since the mice are crosses of inbred founders. Note that other marker-based estimates of the kinship coefficients have been used, and these are based on different derivations of the polygenic covariance matrix. To account for uncertainty in missing genotypes at a given marker, we calculated the expected number of shared alleles.

The final marker-based kinship coefficient estimate was obtained by averaging over markers across the genome, excluding the X chromosome. To avoid “proximal contamination” when assessing evidence for a QTL at a given SNP, this SNP, and all nearby SNPs, should be excluded from estimation of the kinship coefficient (Listgarten et al. 2012; Cheng et al. 2013; Yang et al. 2014). We accomplished this by omitting markers on the same chromosome to estimate relatedness. Therefore, we fitted the mixed model to the data separately for each chromosome and used this mixed model only to quantify support for QTL on that chromosome.

**Covariates:** In all analyses, we included four covariates in the regression model of the phenotype: age (in days) and three binary traits, sex $(1 = \text{male}, 0 = \text{female})$, albino $(1 = \text{white coat color}, 0 = \text{non-white})$, and agouti $(1 = \text{agouti coat color}, 0 = \text{nonagouti})$. We found that coat color confounded some of the phenotype measurements obtained from video tracking; if we did not account for the effect of coat color, for some phenotypes we obtained strong evidence for a QTL that mapped to a region overlapping the Tyr gene on chromosome 7, which is the gene variant for the albino trait (Jackson 1994, 1997). Some phenotype and coat color observations were not recorded for some mice; at most 75 samples in any one phenotype were excluded from QTL mapping due to missing phenotype or covariate values.

A larger proportion of $F_{34}$ mice showed less freezing to context and freezing to cue than $F_2$ mice. However, this difference disappeared if we excluded albino mice, reinforcing the importance of including coat color traits as covariates in the QTL mapping. We observed no systematic differences in distributions of other phenotypes from the $F_2$ and $F_{34}$ crosses.

**Transformation of quantitative traits:** We mapped QTL for six phenotypes: the three measurements recorded in open field testing (distance traveled in the periphery, distance traveled in the center, and percentage of time in the center of the arena) and the three measurements from our conditioned fear tests (pretraining freezing, freezing to context, and freezing to cue). Four of these phenotypes (percentage of time in center, pretraining freezing, freezing to context, and freezing to cue) are proportions between 0 and 1. To obtain numbers on the real line, and to admit a normal model for these phenotypes, we transformed the proportions to the log-odds scale using the logit function, $\text{logit}(x) = \log(x/(1-x))$. To avoid extreme values after the transformation, any proportions $<0.01$ or $>0.99$ were fixed to 0.01 and 0.99, respectively. Observed quantiles of the transformed phenotypes, separately in the $F_2$ and $F_{34}$ crosses, after removing linear effects of the covariates, closely matched expected quantiles under the normal distribution (results not shown), suggesting that the normal distribution was a good fit for these phenotypes.

**Determining significance of genotype–phenotype associations:** To calculate significance thresholds for the LOD scores, we must first obtain the distribution of this test statistic under the null distribution. This distribution is commonly estimated by permuting the phenotype values relative to the genotypes (Broman and Sen 2009). However, this approach will not preserve the covariance structure in the samples that is due to unequal relatedness of the individuals and therefore may lead to inadequately stringent significance thresholds and inflated type 1 error rates (Abney et al. 2002; Zou et al. 2005; Aulchenko et al. 2007; Cheng et al. 2010; Cheng and Palmer 2013). A simple alternative to this approach would be to use a Bonferroni correction of the $P$-values calculated from the test statistics. But this would lead to overly stringent significance thresholds because it would ignore correlations between the markers—that is, it would ignore the fact that the association tests are not independent.

Despite these concerns with the standard permutation test (that assumes independence of the samples), we have reason to expect that this permutation test will still provide...
reasonably accurate significance thresholds because (1) we do not observe systematic population stratification in the AIL (e.g., clusters of samples that are more related to each other than they are to other mice) and (2) no individuals are closely related to each other (e.g., no mice from the same litter). Careful inspection of our marker-based estimates of kinship coefficients confirms that expectations 1 and 2 do not apply to our AIL. Therefore, we performed permutation-based tests using R/qtl, ignoring unequal relatedness of the individuals (note that the tests for association still used the full-relatedness model). To obtain an estimate for the significance threshold at the 100(1 − α)th percentile of the test statistic, with α = 0.05, we recorded the maximum LOD score for each permutation replicate and defined the significance threshold to be the empirical percentile obtained from 1000 maximum LOD scores, following the recommendation of Churchill and Doerge (1994).

To provide an independent validation of this significance threshold, which is based on the assumption that the samples are exchangeable, we compared it against a more computationally intensive permutation test that accounts for relatedness among individuals (Cheng et al. 2010; Cheng and Palmer 2013), similar to the permutation test proposed by Aulchenko et al. (2007). This permutation procedure differs from the standard method in two key ways: (1) it estimates parameters of a “null” model that includes the polygenic effect, in which kinship coefficients are calculated using the markers, and (2) instead of permuting the phenotypes, it permutes the genotypes, which preserves the relationship between phenotype and polygenic effect. (Note that the accuracy of this permutation procedure hinges on how well the polygenic covariance matrix captures the true covariance structure in the phenotype, which is, of course, unknown.) Following our procedure described above, we fitted a separate mixed model for each chromosome during the permutation tests to avoid the problem of proximal contamination. In Results, we show empirically that this more sophisticated permutation-based test produced significance thresholds that closely corresponded to the thresholds obtained from the much simpler permutation procedure using R/qtl.

QTL regions: We used a 1.5-LOD support interval to approximate the confidence interval for the location of the QTL. This is a slightly smaller interval than the suggested 1.8-LOD interval based on simulations in two intercrosses, although in practice the best interval for each QTL depends on a number of factors, including the QTL effect size (Manichaikul et al. 2006). The main reason we chose this interval was to be consistent with our previous studies using AIL mice (e.g., Parker et al. 2011; Bartnikas et al. 2012).

Assessment of multiple QTL: For several phenotypes, the initial genome-wide scan indicated appreciable support for QTL at multiple locations on the same chromosome. However, testing each marker one at a time cannot indicate whether there is support for multiple QTL on a chromosome. To address this question, we included the SNP with the highest LOD score as a covariate, and then we recalculated the LOD scores for all other SNPs on the same chromosome. We repeated this procedure for each of the 30 QTL regions identified in the initial genome-wide scans for the six phenotypes.

X chromosome: QTL mapping must be performed differently in the X chromosome due to differences in males and females (Broman and Sen 2009; Wise et al. 2013) and due to recombination frequencies that are specific to the X chromosome (Broman 2012). Therefore, a proper analysis of the X chromosome that corrects for confounding due to familial relationships is beyond the capabilities of QTLRel. Nonetheless, we used QTLRel to obtain rough estimates of support for QTL to investigate whether there was a suggestion that markers might warrant further investigation. To adapt the mixed model to analyze the X chromosome, we included sex as an interactive covariate—that is, we included separate additive and dominance effects for males and females. We then estimated the null distribution of the test statistic using the same model (without marker effects), using 1000 permutation replicates as before. Note that this model has one too many degrees of freedom (there is no need to include both additive and dominance terms in hemizygous males since there are only two possible genotypes), so our analysis was overly conservative, and one could improve on our X chromosome analysis, for example, following the approach of Pan et al. (2007). On the X chromosome, we computed genotype probabilities for missing genotypes separately for females and hemizygous males.

Bioinformatics databases: To identify genes containing non-synonymous coding SNPs within 1.5-LOD support intervals, we used sequence data for LG/J and SM/J inbred strains generously provided by James Cheverud and Heather Lawson from the Genome Sequencing Center at Washington University [http://genome.wustl.edu (Norgard et al. 2011)].

The full code and data reproducing the steps of the analyses are available for download at http://github.com/pcarbo/lgsmfear.

Results

QTL mapping for open field and conditioned fear phenotypes

We performed a genome-wide association analysis for six traits: (1) distance traveled in the periphery of the open field, (2) distance traveled in the center of the open field, (3) percentage of time spent in the center of the open field, (4) pretraining freezing, (5) freezing to context, and (6) freezing to cue. To assess support for genotype–phenotype correlations at 4601 candidate SNPs, we used a linear mixed model that accounted for possible confounding due to differences in genetic sharing, in which levels of sharing were estimated from the genotype data. We assessed evidence for QTL separately in the LG/J × SM/J F2 intercross, in
the F\textsubscript{34} advanced intercross, and in the combined sample (Figure 1, Figure 2, and Table 1).

In the analysis of the combined sample, we identified a total of 24 QTL exceeding significance thresholds (we discuss the significance thresholds in greater detail below). Specifically, we identified 12 QTL for OF phenotypes and 12 QTL for CF phenotypes, and we identified at least 1 QTL for each of the six behavioral traits.

Our analysis did not yield any compelling QTL on the X chromosome; the largest LOD score, 4.6, for the phenotype “percentage of time spent in center,” was considerably below the significance threshold, 5.7 (at \( \alpha = 0.05 \)). Since sex-linked chromosomes require a separate treatment from that of autosomal chromosomes, X chromosome results are not shown in Table 1, Figure 1, and Figure 2. (Note that our analysis of the X chromosome was probably overly conservative, and an improved analysis of the X chromosome correcting for relatedness could yield stronger support for QTL; see Materials and Methods for more details.)

The initial genome-wide scan (Figure 1 and Figure 2) identified multiple distinct regions located on the same chromosome containing significant QTL (Table 1). In all cases except one—chromosome 2 for freezing to cue—support for a second QTL fell short of our significance thresholds once we conditioned on the estimated effect of the peak SNP from the same chromosome. Table 1 gives two regions on chromosome 17 that each show strong evidence for containing a QTL for freezing to context, although we have sufficient support for a QTL in only one of the two loci.

Several phenotypes were associated with regions that appear to overlap each other (Table 1). However, the QTL mapping did not yield sufficient resolution to determine whether the overlapping QTL regions highlight the same gene or genes. Below, we examine some of the more interesting overlapping QTL.

Most of the QTL with the strongest support from the combined sample were also identified using the F\textsubscript{2} sample alone. One exception was the QTL for freezing to cue on chromosome 2. This locus showed no appreciable association signal in the F\textsubscript{2} sample, but strong support from the F\textsubscript{34} data.

In the F\textsubscript{2} analyses, each of the QTL regions (based on 1.5-LOD support intervals) covered a large portion of the chromosome, owing to limited recombination in F\textsubscript{2} crosses. Compared to the QTL regions in the F\textsubscript{2} cross, the QTL were much narrower in the F\textsubscript{34} alone. However, few SNPs reached the threshold for significance at \( \alpha = 0.05 \), owing to a loss of power due to many more candidate haplotypes. By comparison, in the combined analysis, we identified 24 QTL with narrow 1.5-LOD support intervals, ranging in size from 1.7 Mb to 29.4 Mb and with a median interval length of 4.3 Mb. Of these, 13 QTL had support intervals <5 Mb in length. The number of annotated genes within these intervals ranged from 10 to 236, with a median of 71 genes.

As expected, support for QTL in the F\textsubscript{2} sample for the most part did not change appreciably when we accounted for possible confounding due to relatedness; F\textsubscript{2} crosses are full sibs with respect to one another and hence are expected to exhibit roughly the same amount of genetic sharing. Nonetheless, there were several cases where support for a QTL increased notably in the F\textsubscript{2} cross after accounting for varying relatedness and occasionally yielded a significant QTL only after correcting for relatedness. This comparison was not systematic, nor did we experience a uniform increase in support across the genome after correcting for relatedness. Nonetheless, this hints at the potential gains in applying mixed models to conventional intercrosses—a topic that we plan to explore in greater depth in subsequent work.

Since we had kept track of the complete pedigree for the F\textsubscript{34} mice, we were able to compare the QTL mapping results, correcting for relatedness using (1) marker-based estimates of the kinship coefficients and (2) pedigree-based estimates. Both analyses showed broad agreement; compare the dark blue and light red lines in Figure 1, J–L, and Figure 2, M–O. In several regions, the association signal was considerably stronger in the analysis using marker-based estimates. A notable exception to this tendency was the QTL on chromosome 4 for freezing to cue. These results are consistent with a more systematic comparison of pedigree- and marker-based estimates of relatedness for QTL mapping in simulated populations (Cheng et al. 2013).

Comparison of the pedigree- and marker-based estimates of the kinship coefficients also offers the opportunity to point out some features of the marker-based estimates. Figure 3 shows the distribution of pairwise relatedness coefficients \( R_{ij} = 2\phi_{ij} \) in the F\textsubscript{2} and F\textsubscript{34} crosses. First, using the marker data we were able to predict that some F\textsubscript{2} mice share a greater proportion of their genome than others (Figure 3B), whereas pedigree-based estimates of sharing are all identical, as the relationship to the inbred founders is the same for all F\textsubscript{2} crosses. (For any F\textsubscript{2} cross, pairwise relatedness is \( R_{ij} = 3/2 \) and \( R_{ij} = 1 \), for \( i \neq j \); see Lange 2002.) By comparison, using the marker data we predict much less variation in genetic sharing between pairs of F\textsubscript{34} generation mice (compare Figure 3B with Figure 3D). This was expected because of the greater number of accumulated crossovers in the F\textsubscript{34} mice. Second, we observed that the marker-based estimates were for the most part unbiased; differences in the marker- and pedigree-based estimates were centered near zero (see Figure 3, A, B, E, and F). Even though the genetic sharing estimates in F\textsubscript{34} ALL mice agreed on average, the estimates did not agree in a large proportion of the samples—the scatterplot (Figure 3G) illustrates this point in detail. The larger discrepancies in the sharing estimates may explain the differences observed in the QTL mapping (Figure 1 and Figure 2). Finally, the marker data allowed us to estimate sharing for F\textsubscript{2}–F\textsubscript{34} pairs of mice (Figure 3H), whereas these calculations are very complex using the pedigree data (and would not have been meaningful in this case because the F\textsubscript{2} are not the direct progenitors of the F\textsubscript{34}).

To assess significance of the LOD scores in our genome-wide scans, we applied a simple permutation test (see Materials and Methods for details). Although this simple permutation test can lead to inadequately stringent significance thresholds when phenotype samples are correlated (Abney et al. 2002; Zou et al. 2005; Aulchenko et al. 2007; Cheng et al. 2010;
Cheng and Palmer 2013), as they are here, we have assessed empirically in our AIL that the distribution of the test statistic under the null is similar regardless of whether the permutations account for relatedness (Figure 1 and Figure 2). One possible explanation why the standard permutation test seems to adequately control for type 1 error in our AIL sample is that we have carefully avoided including highly related mice—particularly, no mice share a common grandparent—unlike our previous AIL studies (e.g., Cheng et al. 2010; Cheng and Palmer 2013).

In summary, the combined analysis identified more QTL than either the F2 or the F34 samples alone (due in part to a larger sample size). Importantly, the combined analysis located the QTL within narrower regions compared to those identified from the F2 mice alone.

Sequence analysis
We examined our 1.5-LOD support intervals for the presence of “consequential” SNPs that had the potential to directly alter proteins (i.e., nonsynonymous coding, stop-gain, stop-lost, frameshift, and splice sites). The number of genes with consequential SNPs in significant QTL intervals ranged from 1 to 59, with a median of 12 genes (Table S1).
Figure 2 Conditioned fear and QTL genome-wide map of conditioned fear traits in LG/J × SM/J F2, F34, and combined cohorts. A–F summarize conditioned fear (CF) testing in mice. A 3-day procedure was used to phenotype each subject. On day 1, a baseline indicator of freezing tendency, “pretraining freezing,” was measured in F2 (A) and F34 (D) mice from 30 to 180 sec, after which mice were exposed to two 30-sec tones (indicated by gray bars, labeled T1 and T2) that coterminated with a 2-sec, 0.5-mA foot shock (indicated by arrows, labeled S1 and S2). On day 2, freezing to context was measured in F2 (B) and F34 (E) mice from 30 to 180 sec. On day 3, freezing to cue was measured in F2 (C) and F34 (F); time spent freezing to each tone was averaged to obtain the freezing to cue variable (180–210 sec + 240–270 sec). Each data point represents average percentage of freezing calculated across the 30-sec time bin; error bars represent standard error. G–O show the genome-wide association signal from F2 samples alone (G–I), from the F34 samples alone (J–L), and from the combined F2 + F34 cohort (M–O). See Figure 1 legend for a detailed explanation of the panels.
<table>
<thead>
<tr>
<th>Chr</th>
<th>Peak LOD</th>
<th>Peak SNP</th>
<th>1.5-LOD interval (Mb)</th>
<th>QTL width</th>
<th>rs start</th>
<th>rs end</th>
<th>No. genes</th>
<th>Genes with NS SNPs</th>
<th>LG/J homozygote mean</th>
<th>heterozygote mean</th>
<th>SMJ homozygote mean</th>
<th>PVE</th>
<th>Source for QTL previously implicated in fear/anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6.20</td>
<td>rs27100797</td>
<td>9.594–26.988</td>
<td>17.39</td>
<td>rs33088454</td>
<td>rs13476394</td>
<td>177</td>
<td>6</td>
<td>2614.77</td>
<td>2843.57</td>
<td>2780.70</td>
<td>0.024</td>
<td>Sokoloff et al. (2011)</td>
</tr>
<tr>
<td>4</td>
<td>6.75</td>
<td>rs32242095</td>
<td>30.271–33.911</td>
<td>3.64</td>
<td>rs28717374</td>
<td>rs2790884</td>
<td>15</td>
<td>4</td>
<td>2626.90</td>
<td>2789.62</td>
<td>2939.04</td>
<td>0.026</td>
<td>Sokoloff et al. (2011)</td>
</tr>
<tr>
<td>11</td>
<td>6.54</td>
<td>rs29391262</td>
<td>73.660–81.992</td>
<td>8.33</td>
<td>rs29453763</td>
<td>rs9431896</td>
<td>150</td>
<td>42</td>
<td>2908.44</td>
<td>2836.54</td>
<td>2746.93</td>
<td>0.025</td>
<td>Sokoloff et al. (2011)</td>
</tr>
<tr>
<td>14</td>
<td>4.51</td>
<td>rs6289260</td>
<td>111.373–18.862</td>
<td>7.49</td>
<td>rs3670736</td>
<td>rs30259301</td>
<td>10</td>
<td>1</td>
<td>2818.45</td>
<td>2836.09</td>
<td>2634.62</td>
<td>0.017</td>
<td>Sokoloff et al. (2011)</td>
</tr>
<tr>
<td>17</td>
<td>5.01</td>
<td>rs29501002</td>
<td>25.869–29.205</td>
<td>3.34</td>
<td>rs6303335</td>
<td>rs4911021</td>
<td>78</td>
<td>23</td>
<td>2548.33</td>
<td>2783.25</td>
<td>3002.24</td>
<td>0.019</td>
<td>Sokoloff et al. (2011)</td>
</tr>
<tr>
<td>4</td>
<td>5.11</td>
<td>rs32263942</td>
<td>118.880–121.182</td>
<td>2.30</td>
<td>rs32320402</td>
<td>rs31878303</td>
<td>40</td>
<td>8</td>
<td>596.55</td>
<td>527.09</td>
<td>442.72</td>
<td>0.019</td>
<td>Nakamura et al. (2003)</td>
</tr>
<tr>
<td>10</td>
<td>4.44</td>
<td>rs62636734</td>
<td>127.334–129.668</td>
<td>1.72</td>
<td>rs6157058</td>
<td>rs6236374</td>
<td>78</td>
<td>33</td>
<td>585.54</td>
<td>513.47</td>
<td>508.95</td>
<td>0.017</td>
<td>Jones et al. (1999); Wilson et al. (2011)</td>
</tr>
<tr>
<td>12</td>
<td>5.12</td>
<td>rs6176279</td>
<td>55.962–58.325</td>
<td>2.96</td>
<td>rs29212678</td>
<td>rs2921863</td>
<td>14</td>
<td>4</td>
<td>456.63</td>
<td>520.83</td>
<td>618.88</td>
<td>0.019</td>
<td>Sokoloff et al. (2011)</td>
</tr>
<tr>
<td>4</td>
<td>5.06</td>
<td>rs36318014</td>
<td>104.577–114.064</td>
<td>9.49</td>
<td>rs32431061</td>
<td>NA</td>
<td>83</td>
<td>15</td>
<td>11.93</td>
<td>11.05</td>
<td>10.77</td>
<td>0.019</td>
<td>Nakamura et al. (2003)</td>
</tr>
<tr>
<td>10</td>
<td>4.42</td>
<td>rs29369956</td>
<td>125.235–129.686</td>
<td>3.83</td>
<td>rs29334025</td>
<td>rs623674</td>
<td>104</td>
<td>48</td>
<td>12.15</td>
<td>10.80</td>
<td>10.64</td>
<td>0.016</td>
<td>Sokoloff et al. (2011)</td>
</tr>
<tr>
<td>12</td>
<td>5.12</td>
<td>rs6176279</td>
<td>54.210–58.331</td>
<td>4.12</td>
<td>rs29174777</td>
<td>rs2913440</td>
<td>38</td>
<td>3</td>
<td>11.05</td>
<td>11.15</td>
<td>12.40</td>
<td>0.019</td>
<td>Wilson et al. (2011)</td>
</tr>
<tr>
<td>15</td>
<td>5.46</td>
<td>rs32422569</td>
<td>34.289–38.377</td>
<td>4.09</td>
<td>rs32100330</td>
<td>rs32514574</td>
<td>31</td>
<td>8</td>
<td>10.45</td>
<td>10.80</td>
<td>11.96</td>
<td>0.020</td>
<td>Sokoloff et al. (2011)</td>
</tr>
<tr>
<td>17a</td>
<td>3.99</td>
<td>rs33357930</td>
<td>158.322–161.410</td>
<td>8.83</td>
<td>rs3086080</td>
<td>rs3082730</td>
<td>71</td>
<td>6</td>
<td>29.85</td>
<td>35.89</td>
<td>37.81</td>
<td>0.018</td>
<td>Sokoloff et al. (2011); Wilson et al. (2011); Radcliffe et al. (2000); Balden et al. (1997)</td>
</tr>
</tbody>
</table>

Table 1. QTL for open field and conditioned fear phenotypes

This table lists QTLs for open field activity and conditioned fear phenotypes. The columns provide information such as chromosome (Chr), peak LOD score, peak SNP, 1.5-LOD interval in megabases (Mb), QTL width, rs start and end positions, number of genes within the region, genes with nonsynonymous coding polymorphisms, LG/J homozygote mean, heterozygote mean, SMJ homozygote mean, and PVE. Additionally, it notes the source of QTLs previously implicated in fear/anxiety.

Significance thresholds are defined as the 95th percentile of the test statistic under the null (α = 0.05). LOD scores were obtained by fitting a model to the combined (F₂ + F₄) cohort, using marker-based estimates of pairwise relatedness. Columns from left to right are: (1) chromosome, (2) maximum LOD score, (3) SNP with maximum LOD score, (4) and 5) region most likely containing causal variants and size of region in megabases (Mb) based on a 1.5-LOD support interval, (6) SNPs bounding region, (7) number of genes within region annotated to mouse reference sequence, (8) number of genes within region for which nonsynonymous coding polymorphisms are identified (see Table S1), (9) mean phenotypic value for LG/J homozygotes at peak SNP, (10) LG/J × SMJ heterozygotes, (11) SMJ homozygotes, (12) proportion of variance in phenotype explained by peak SNP, and (13) published findings on fear and anxiety QTL that overlap this region. All SNP information and genomic positions are based on Mouse Genome Assembly 37 (July 2007) in the NCBI Assembly Database.

Support for the other QTL on chromosome 17 falls short of our criteria for significance once we condition on this QTL; our data support at most one QTL on chromosome 17.

SNPs in this region exceed the significance threshold only after conditioning on the effect of SNP rs27453770 on the same chromosome.
point in the scatterplot corresponds to a pair of $F_{34}$ mice. In the $F_2$ cross, all pedigree-based pairwise relatedness estimates are $R_{ij}$, where $i$ and $j$ are different individuals ($i \neq j$) or when $i = j$ (in which case the coefficients correspond to inbreeding levels). These are “realized relatedness” estimates obtained by calculating average allele sharing over all SNPs genotyped in the $F_2$ and $F_{34}$ crosses. Diagonal entries of the relatedness matrix, $R_{ij} = 2R_{ij}$, are shown in A and C, and off-diagonal entries $R_{ij} = 2R_{ij}$, $i \neq j$, are shown in B, D, G, and H. To compare the marker-based estimates to the estimates of allele sharing obtained from the AIL pedigree data, in E and F we show the distribution of differences between the marker-based and pedigree estimates of pairwise relatedness. In G we show a scatterplot providing a more detailed comparison of the marker- and pedigree-based pairwise relatedness estimates $R_{ij}$, $i \neq j$, in which each

### Figure 3

Comparison of marker- and pedigree-based estimates of pairwise relatedness. A–D and H show the extent to which relatedness, or genetic sharing, varies among the AIL mice. More precisely, A–D and H give the observed distributions of estimated pairwise relatedness coefficients, or two times the kinship coefficients, for all pairs of mice $(i,j)$ in which $i$ and $j$ are different individuals ($i \neq j$) or when $i = j$ (in which case the coefficients correspond to inbreeding levels). These are “realized relatedness” estimates obtained by calculating average allele sharing over all SNPs genotyped in the $F_2$ and $F_{34}$ crosses. Diagonal entries of the relatedness matrix, $R_{ij} = 2R_{ij}$, are shown in A and C, and off-diagonal entries $R_{ij} = 2R_{ij}$, $i \neq j$, are shown in B, D, G, and H. To compare the marker-based estimates to the estimates of allele sharing obtained from the AIL pedigree data, in E and F we show the distribution of differences between the marker-based and pedigree estimates of pairwise relatedness. In G we show a scatterplot providing a more detailed comparison of the marker- and pedigree-based pairwise relatedness estimates $R_{ij}$, $i \neq j$, in which each

### Discussion

We investigated the genetic basis of anxiety-like behavior, based on open field and conditioned fear tests, in a combined $F_2$ and $F_{34}$ AIL mouse population. To map QTL across the genome, we analyzed phenotypes using a linear mixed model that accounts for differences in genetic sharing, an important confounding factor in tests for association. Taking this mixed-model approach, we identified a total of 24 QTL affecting six traits related to fear and anxiety-like behavior. Because our study maintained uniform environmental conditions and ensured alleles with high frequency, we expect a limited contribution of gene–environment interactions to variation in these traits.

The integration of the $F_2$ and $F_{34}$ populations provided good power to detect QTL, presumably because of the $F_2$ cross; and good resolution, presumably because of the $F_{34}$ advanced intercross. The median size of a QTL region was 4.3 Mb. The QTL were slightly narrower than QTL intervals we identified in this population for other traits, such as red blood cell parameters [median 1.5-LOD support interval = 4.7 Mb (Bartnikas et al. 2012)] and body weight [median 1.5-LOD support interval = 5.5 Mb (Parker et al. 2011)].

In the 1.5-LOD intervals that contained $\leq 15$ genes with coding SNPs, we searched for candidate genes that could plausibly explain the variation in the traits we measured. Some of the genes within these regions had coding SNPs known to be involved in anxiety-like behavior and/or conditioned fear; examples include the nuclear receptor subfamily 6, group A, member 1 gene [$Nrr6a1$ (Heydendael et al. 2013]), the phospholipase D1 gene [$Pld1$ (Sun et al. 2013)], the cadherin 23 gene [$Cdh23$ (Terracciano et al. 2010, but see Schwander et al. 2009)], the prosaposin gene [$Psap$ (Hovatta et al. 2005; Donner et al. 2008)], and the SLIT and NTRK-like family, member 5 gene [$Slitrk5$ (Shmelkov et al. 2010)]. However, we are cautious about interpreting the functional relevance of nonsynonymous coding SNPs; we do not currently have expression QTL (eQTL) data that could be used to detect heritable regulatory polymorphisms. Thus, the current results are not sufficient to identify specific causal genes.

In two instances, contextual and cued fear mapped to overlapping chromosomal regions on chromosomes 2 and 17 (Figure S2 and Figure S3) despite known differences in their neuroanatomical substrates (Fanselow and Ledoux 1999; Jovanovic and Ressler 2010). This may indicate the presence of alleles that influence both traits; alternatively, it could be due to different alleles that are located close to each other in the genome. We (Ponder et al. 2007a,b; Sokoloff et al. 2011; Parker et al. 2012b) and others have reported similar results; Talbot et al. (2003) reported that contextual and cued fear were highly correlated ($r = 0.63$) in heterogeneous stock mice, and selection for freezing to context has been shown to cause coincident changes in freezing to cue (Radcliffe et al. 2000; Ponder et al. 2007a). Thus, it is likely that contextual and cued fear are modulated by some of the same alleles, but further studies are needed to definitively identify the alleles that give rise to such correlations.

One of the underlying motivations for this study was the belief that a subset of alleles would pleiotropically influence both conditioned fear and anxiety-like behavior in the open field. Previous studies by our laboratory (Ponder et al. 2007a,b; Sokoloff et al. 2011) and other groups (Lopez-Aumatell
et al. 2009) have indicated a shared genetic control of these two traits. The overlap between the QTL on chromosome 17 for conditioned fear (freezing to context and freezing to cue) and distance traveled in the periphery (Figure S3) provides some additional support for a shared genetic basis. Interestingly, we observed QTL on chromosome 2 for freezing to context and freezing to cue that were immediately adjacent to, but did not overlap with, the chromosome 2 QTL for distance traveled in the periphery (Figure S2). And in one case, the QTL on chromosome 4 for pretraining freezing showed modest overlap with the chromosome 4 QTL for distance traveled in the center of the arena (Figure S4). However, pretraining freezing (freezing before administration of any tones or shocks) is more indicative of unlearned anxiety rather than conditioned fear. The absence of regions common to several traits may indicate that in this population the genetic origin of fear and anxiety-like behavior is relatively distinct, or it may reflect a lack of power to detect QTL, so we cannot interpret this result as implying that genes in these regions do not jointly affect these traits. An integrated analysis of these traits [e.g., building on multivariate mapping approaches (Stephens 2013)] may yield an improved understanding of how QTL modulate these anxiety-like behaviors and represents an important research direction.

Traditionally, F2 intercrosses are used to identify QTL underlying phenotypic variation, and fine-mapping is carried out as a second step, using congenic strains. This time- and labor-intensive effort to identify specific genes is often derailed by the discovery that a single QTL of large effect is in fact caused by multiple loci of small effect located in the same chromosomal region (Legare et al. 2000; Mott et al. 2000; Cheng et al. 2010; Shao et al. 2010; Parker et al. 2013). An AIL is an improvement over these traditional methods because it merges identification and fine-mapping into a single step, which can often discriminate between loci that are due to single vs. multiple alleles (Darvasi and Soller 1995). The trade-off is that the power to detect QTL in AILs is often lower than in F2 populations. This is because the AIL mice experience greater numbers of crossover events than the F2 mice, so more tests are performed, and a corresponding higher threshold is needed to control for false positives. For example, a QTL was observed on chromosome 1 for freezing to context in the F2 intercross, but in the F34 AIL it appeared to split into two smaller regions, neither of which had a LOD score exceeding significance. On the other hand, we observed a highly significant QTL peak in the F34 AIL for freezing to cue that was not originally seen in the F2 intercross and was subsequently supported in the integrated analysis. While the advantages of an integrated analysis of an AIL have been shown in previous work, this study provides further support for the benefits of this approach and can serve as a prototype for how to identify QTL in AILs using a mixed-model approach that accounts for relatedness, using marker data. (We have made available the code and data used to implement the steps of our analysis at http://github.com/pcarbo/lgsmfear.)

Our main methodological contribution was to show that we can use marker data in the place of a pedigree to infer familial relationships in an AIL. We found that mixed models using either pedigree- or marker-based estimates of relatedness showed broad agreement in the QTL mapping. While there has been a considerable amount of work demonstrating the benefits of marker-based estimates of “realized relatedness” to control for confounding due to population structure or due to familial relationships (Yu et al. 2006; Kang et al. 2008; Lippert et al. 2011; Listgarten et al. 2012; Zhou et al. 2013), there has been little work on demonstrating these benefits in AILs. In this article, we did not aim for a systematic comparison of approaches to correct for relatedness (for an empirical comparison in simulated data sets, see Cheng et al. 2013). Our results nonetheless suggest that it is better to use marker data to correct for relatedness, even when the full pedigree is available, provided genetic variation is ascertained at sufficient resolution throughout the genome. The marker-based estimates often yield more precise estimates of genetic sharing and are usually less costly to obtain.

Our study has several limitations. First, because we have used a cross between two inbred strains, we are studying the alleles that segregate between them and not the total number of alleles that segregate among other laboratory strains or wild mice. For example, we observed little overlap of QTL for conditioned fear between an F8 AIL derived from C57BL/6J × DBA/2J mice (Parker et al. 2012b) and the LG/J × SM/J AIL in the present study. Nonetheless, we did observe some overlap in the QTL we identified in our population with anxiety and fear-related QTL identified in other populations of mice (Table 1), consistent with the possibility that different two-strain combinations may segregate the same alleles. It is possible that some of the QTL identified in our study are the same as those identified by other researchers; one advantage of our study is that we have mapped QTL with greater resolution than in previous studies.

Our approach has dramatically increased the mapping resolution compared to that in more conventional mapping populations. To further reduce the number of candidate genes within our intervals, we focused on genes with coding polymorphisms. However, it is important to note that the polymorphisms underlying the observed trait variance may be due to differences in gene expression, rather than alterations in gene function. For example, some QTL may be explained by SNPs that modulate the recruitment of proteins involved in regulation of gene transcription; studies of complex human traits have shown that a large fraction of the variants underlying these traits coincide with DNA sequences related to gene regulation (Nicolae et al. 2010; Schaub et al. 2012). Availability of genome-wide eQTL data in the LG/J and SM/J strains may help pinpoint the genes underlying these QTL.

Acknowledgments
We thank James Cheverud and Heather Lawson for generously providing the F33 mice used to create the F34 and the
sequence data in the LG/J and SM/J mice. We also acknowledge Michaelanne Munoz, Tanya Cebollero, and Jackie Lim for help with husbandry and behavioral testing; Kaitlin Samocha for pedigree reconstruction; Karl Broman for advice on dealing with QTL mapping on the X chromosome; Xiang Zhou for technical help on the QTL mapping; Pei-Ting Wu for data entry; and Ryan Cheng and Andrew Skol for helping to develop QTLRel and calling of genotypes used in this study. Finally, we thank the reviewers for their helpful feedback. This work was supported by National Institutes of Health grants 5R01MH079103 (to A.A.P.) and R01HG002899 (to M.A.). P.C. was supported by a cross-disciplinary postdoctoral fellowship from the Human Frontiers Science Program.

Literature Cited
High-Resolution Genetic Mapping of Complex Traits from a Combined Analysis of F2 and Advanced Intercross Mice

Clarissa C. Parker, Peter Carbonetto, Greta Sokoloff, Yeonhee J. Park, Mark Abney, and Abraham A. Palmer
Table S1  Candidate non-synonymous coding SNPs for Open Field and Conditioned Fear QTLs. Table columns from left to right are: (1) chromosome; (2) base pair position of SNP, and strand; (3) gene; (4) gene symbol; (5) total number of nonsynonymous SNPs identified in gene. Available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.167056/-/DC1

Figure S1  Comparison of reference genetic map and genetic distances estimated from F2 genotype data. Each point in scatterplot corresponds to a single SNP on the X chromosome (orange) or an autosomal chromosome (black). Intermarker distances are shown in centiMorgans (cM).
Figure S2  Overlapping QTL region on chromosome 2 for freezing to context and freezing to cue, and adjacent QTL region on chromosome 2 for distance traveled in the periphery. LOD scores quantify support for association at SNPs, accounting for differences in genetic sharing using marker-based estimates of pairwise relatedness. Note that LOD scores for freezing to cue are obtained by conditioning on effect of peak SNP (rs27453770) on the same chromosome. Horizontal red lines indicate 1.5-LOD support intervals. The dashed orange line shows the significance threshold at $\alpha = 0.05$ obtained using the simple permutation test.
Figure S3  Overlapping QTL region on chromosome 17 for freezing to context, freezing to cue and distance traveled in the periphery. LOD scores quantify support for association at SNPs, accounting for differences in genetic sharing using marker-based estimates of pairwise relatedness. Red lines represent 1.5-LOD support intervals, and the dashed orange line shows the significance threshold (at α = 0.05) obtained using the simple permutation test.
Figure S4  Overlapping QTL region on chromosome 4 for distance traveled in the center and pre-training freezing. LOD scores quantify support for association at SNPs, accounting for differences in genetic sharing using marker-based estimates of pairwise relatedness. Red lines represent 1.5-LOD support intervals, and the dashed orange line shows the significance threshold (at $\alpha = 0.05$) obtained using the simple permutation test.